

Slide 1 of 7

Title Slide: Getting Started with the ECLS-B Data

Slide 2 of 7

This module provides users an overview of the data collected through the ECLS-B that are available for researchers to analyze. The overview is intended to help users begin to develop a conceptual understanding of the micro-level data file containing the data produced from the study. This module also describes the resources that are available to learn more about the study, the data, and the data file. Information presented in this module will be helpful in understanding some of the more detailed information presented in subsequent ECLS-B modules. For this reason, users who are planning to proceed through the subsequent ECLS-B modules and use ECLS-B data for analytic purposes are strongly encouraged to complete this module first.

Slide 3 of 7

Information pertaining to the ECLS-B children was collected at and about multiple time points in the children's lives -- from birth through kindergarten entry. Information was collected from multiple sources using a variety of data collection methods.

Data from all of these time points, sources of information, and data collection components are available in the ECLS-B data file. More specifically, the data file includes variables providing information obtained from the vast majority of questions asked in the parent interviews, the early care and education provider telephone interviews, and the hard-copy teacher questionnaires. All of these data collection instruments are available on-line and can be accessed by clicking the underlined screen text. Some information obtained from the birth certificate was taken from the National Center for Health Statistics natality files and is included in the ECLS-B data file. Information about the schools children attended in kindergarten was obtained from NCES's school universe files, the Common Core of Data and the Private School Survey. This school information is included in an appendix on the DVD containing the other ECLS-B data and needs to be merged into the main ECLS-B data file.

Information collected through data collection instruments other than the interviews and questionnaires is provided in various ways on the ECLS-B data file, the most common being a variable or series of variables representing overall or general performance on a group of items, as opposed to responses to single or individual items. The data file does include item-level data for some of the child assessments, for example the NCATS and the Two Bags Task. However, item-level data cannot be provided for the direct cognitive assessments due to copyright restrictions, so assessment scores representing children's performance were calculated and provided on the data file for researchers. The audiotapes of children telling stories and videotapes of parent-child interactions are not provided to researchers for privacy reasons. Trained coders reviewed the recordings and evaluated the children's retelling of stories and parent-child interactions

on a number of quality indicators. The codes assigned on these quality indicators are provided in the data file.

The next two modules in this series describe the data collected through the ECLS-B. The first module, 'Child Assessment Data,' describes the assessments that were fielded in the ECLS-B in more detail and provides information about the assessment scores that are available for analysis. The codes assigned on the quality indicators used in the child care observation are also provided for analysis. The second module, 'Early Care and Education Data' provides more detail about characteristics of children's early care and education settings that were evaluated in the observation component of the study.

Slide 4 of 7

In addition to information that was obtained directly from one or more respondents in the study or indirectly through observation, there are several other categories of variables available on the data file: administrative variables, composite variables, sampling variables, and weighting variables.

There is a wealth of information about study respondents and different aspects of data collection that is compiled and maintained in administrative records throughout the course of the study. This information typically is not collected from respondents through the data collection instruments. Some administrative information that may be useful for analyses is provided in the data file. Examples of such variables include the ZIP code of the child's residence, the child's age at the time of assessment, and an assessor identification number.

Composite variables are variables that were created using information from two or more variables, two or more sources, or both. Sometimes the composite variables are derived using information that is not available to the researcher in the data file, for example sensitive administrative data that could not be provided for data confidentiality reasons. Detailed descriptions of all the composite variables can be found in the data file user's manuals that are provided on the DVD with the ECLS-B data.

There is a small set of variables on the data file that provide information about how the ECLS-B children were categorized for sampling purposes, more specifically how cases were treated for oversampling. These variables are provided for researchers specifically interested in methodological issues associated with sampling and should not be used for any other kind of analysis.

Lastly, the data file contains many variables that are used in the process of weighting and standard error adjustment. These issues are discussed at a general level in the common modules. The 'ECLS-B Weights, Variance, and Missing Data Module' has information about weighting and standard error adjustment specific to the ECLS-B.

Slide 5 of 7

All of the variables in the ECLS-B data file are named using a standard set of conventions. Although this information is data-file specific, it will be useful to have a basic understanding of these naming conventions as you proceed through the ECLS-B modules, because some of the discussion references variables by their names.

The first character of a variable's name indicates the source of the data. For example, variables that pertain to questions in the parent interview start with the letter 'P', while those from the child assessments begin with the letter 'C'. Composite variables begin with the letter 'X' or 'Y'.

The second character of the variable name indicates the round of data collection, with the numbers 1 through 5 corresponding to a particular round, as shown here.

For example, the first two characters of the variable P3READBO indicate that the variable pertains to a question asked in the round 3, or preschool, parent interview.

The first two characters of the variable C1EXPLRE indicate that the variable pertains to the round 1, or 9-month, child assessment.

At the bottom of the slide, there is a link to a handout that provides a comprehensive list of variable naming conventions for all rounds of collection. You may find it helpful to print this handout and use it as a reference as you proceed through the other modules.

Slide 6 of 7

These modules have been designed to help you understand and analyze ECLS-B data. While completing these modules will provide a solid foundation of knowledge, the modules alone do not provide sufficient detail to ensure successful analyses that reflect a comprehensive understanding of the data and what they mean.

First and foremost anyone interested in using the data should read through the user's manuals and other technical reports, which provide a wealth of information about how the different measures used in the study were developed, what they indicate, and how data from the measures are reported or provided in the data file.

In particular each user's manual has a section called data considerations, caveats, anomalies, and errata that users should familiarize themselves with once they know which variables they are interested in using. All data files undergo a thorough review before they are released to the public. During this process, NCES finds issues in the data that users should know about, for example problems due to administration errors that cannot be fixed, or unusual patterns in the data that may or may not be explainable.

The user's manuals are only available to analysts who are granted a restricted-use license to use the data. However, there are many technical reports that are publicly available on the ECLS-B website. Also, users may want to read through the substantive

Getting Started with the ECLS-B Data

or topical reports that use ECLS-B data to see what kind of information is available and what kinds of analyses can be conducted with the data.

For each round of collection, NCES released a *First Look* report intended to feature data collected at that round. In addition, the *First Look* reports contain more in-depth information about the study design for that round.

It is also important for users to review the data collection instruments that are available online so that they know exactly how questions were asked and understand the groups of people to or about whom questions were asked. As you review the instruments you should pay special attention to the paths through the questionnaire and the skip patterns.

Lastly, it is always a good idea to check the ECLS website occasionally for information about new data releases or issues that have been found in the data of which people should be aware.

Slide 7 of 7

This module has provided you with an overview of the data collected through the ECLS-B that are available for researchers to analyze. This module also described the resources that are available to learn more about the study, the data, and the data file. The module's objectives and the resources provided throughout the module are summarized here for your reference.

You may now proceed to the next module in the series, or click the exit button to return to the landing page.